REMARKS

Claim 2 is objected to because of typographical error of "or" in place of "for". Applicant has cancelled Claims 1 and 2. Claim 3 is placed in independent form. Claim 5 is amended to delete the second "using". Claim 6 is amended to change "bbase" to "base". Claim 9 is cancelled. Claim 4 is rejected as being indefinite since it contains two claims – one dependent on claim 3 and the other dependent on itself as dependent on claim 4. Claim 4 is cancelled.

The specification is amended to aid the understanding on what was disclosed in the original specification. It is believed that the original specification supports the amendments and the substitute specification contains no new matter.

Claims 1-3 and 5 are rejected under 35 U.S.C. 102 (b) as being anticipated by Neumeyer et al. (U.S. patent No. 6,226,611; hereinafter Neumeyer.

Speech recognition devices are typically deployed in different acoustic environments such as speech signal produced by male speakers, female speakers, in office environments or in noisy environments. The typical way of dealing with multiple environments is to train multiple HMM model sets, for example we could train separate male and female HMM model sets since the sounds or models for male speakers and female speakers are different. For a given sentence grammar, if we have M sets of HMMs which represent M different environments, a speech recognizer is required to decode M sets of HMMs each of which models a specific acoustic environment. The requirement for M-sets of sentence networks makes the recognition device more costly and requires much more memory. Applicant describes and claims a new recognition method which needs only to represent the structure of one out of the M sub-networks and gives the same performance by

using a generic speaker independent grammar network composed of base symbols to produce a virtual expanded network of virtual expanded symbols representing the virtual expanded network of HMM sets where the pronunciation of each symbol is specified by a set of HMM states. The new recognizer builds recognition paths defined on the expanded symbols which are defined through a conversion function that gives the base symbol of any expanded symbols, and vice versa.

> Applicant's claim 3 calls for "A speech recognizer for decoding multiple HMM sets using a generic base sentence network comprising: means for decoding HMM sets using the generic base sentence network and a recognizer recognizing speech using said decoded multiple HMM sets wherein the means for decoding includes means for building recognition paths defined on expanded symbols and accessing said network using base symbols through a conversion function that gives the base symbol of any expanded symbols, and vice versa."

Neumeyer teaches a method and system for automatic text-independent grading of pronunciation by a student for language instruction. Computer-aided language instruction systems exercise the listening and reading comprehension skills of language students. In particular, the subject is for a computer-based language instruction system to evaluate the quality of the students' pronunciation. The method and system of the reference assesses the pronunciation quality of an arbitrary speech utterance based on one or more metrics on the utterance, including acoustic unit duration and a posterior-probability –based evaluation. The examiner references for Claim 1, col. 4, lines 48-55, col. 9, lines 25-35 and col. 10, lines 30-45. Neumeyer and the sections that the examiner references in his rejection just discuss existing HMM technology and how it is used in his patent. Nowhere does Neumeyer teach that the HMM speech recognizer can use a generic base sentence grammar to recognize multiple HMM sets (please note the plural) so that the recognizer can determine which set

yields the best likelihood recognition result. In applicant's invention the preferred embodiment of the enumeration of the sets is male and female (though there could be any number or enumeration of sets, such as child, dialect, foreign accent, etc.) Applicant's recognizer will return a recognition result based on the additional constraint that the result came from one and only one of the possible HMM model sets. The novelty is that this is done using a generic base sentence grammar network that does not contain any information about the enumeration sets. Claim 5 of Neumeyer makes it clear that our invention is not taught by calling for "...computing a path through a set of trained hidden Markov models...". Note that his claim 5 specifically uses the singular "set", indicating that the HMM recognizer does not decode the recognition result to be specific to one of a plurality of HMM sets.

Claim 10 dependent on claim 3 is deemed allowable over Neumeyer for at least the same reasons as Claim 3. Claim 10 further calls for "the extensions are implemented in calculating HMM deltas in the processing steps get-offsets and get-true–symbols which interface between the single sentence network object and the multiple environment HMM sets".

Claim 5, as amended, calls for "A speech recognition search method for decoding multiple HMM sets using a generic base sentence network comprising: providing a generic grammar, providing expanded symbols representing a network of expanded sets and building recognition paths defined by the expanded symbols and accessing the generic base network using base symbols through a proper conversion function that gives the base symbol of any expanded symbols, and vice versa." This speech recognition search method is not taught or suggested in the reference for the reasons discussed above in connection with Claim 3. Claims 6-9 are rejected under 35 U.S.C. 102(b) as being anticipated by Naylor et al. (U.S. Patent No. 5,806,034; hereinafter Naylor).

Claim 11 dependent on claim 5 is deemed allowable for at least the same reasons as claim 5.

Claim 6 calls for "A method of speech recognition for decoding multiple HMM sets using a generic base sentence network comprising the steps of: providing a generic network containing base symbols; a plurality of sets of HMMs where each set of HMMs corresponds to a single environmental factor such as for male and female; each said set of HMMs enumerated in terms of expanded symbols which map to the generic network base symbols; accessing said generic network using said base symbols through a conversion function that gives base symbols for expanded symbols to therefore decode multiple HMM sets using a generic sentence grammar and using said HMM sets to recognize incoming speech."

Naylor is not applicable and irrelevant since it does not teach an HMM recognizer that decodes a plurality of HMM sets and returns the recognition result. That is, it provides the best recognition result from a single HMM set.

The examiner states that Naylor teaches a method of speech recognition comprising the step of providing a generic network containing base symbols. It is true that Naylor discusses an HMM speech recognizer. That is about all that can be said to be in common with our application. The examiner refers to Fig. 2 to say that Naylor teaches a "generic network containing base symbols" Fig. 2 is actually a flow chart of HMM model training, and has nothing to do with a sentence grammar network used during recognition. The reference is not correct and is not applicable. The examiner references " a single set of HMMs for male and female (as training HMMs for male and female-col. 6, lines 15-25). It is not seen how this can be extracted from col. 6, lines 15-25 since this reference is discussing about raw data collection , and this does not mention how the raw data is used to train HMM models. Applicant has amended Claim 6 to more clearly present applicant's invention.

The examiner references Fig. 4 of the reference to teach "building recognition paths defined on virtual symbols corresponding to base symbols (as building paths using base HMMs in Fig. 4). Fig. 4 of Naylor has nothing to do with a generic sentence grammar network and a mapping of base to virtual symbols. Fig. 4 is just an overall flowchart of a recognizer.

The examiner then makes the following assertions:

"accessing said generic network using said base symbols through conversion function that gives base symbols for virtual symbols (as building upon the base with new model information at each node- Figs.5-7); to therefore decode multiple HMM sets using a single sentence grammar and using said HMM sets to recognize incoming speech performing the recognition – Col. 8, lines 45-52; using grammar sentence models- Co. 7, lines 49-55)."

Figs. 5-7 are just steps any HMM recognizer takes. There is nothing in the figures or the referenced text that correspond to a speech recognizer using multiple HMM sets which map to a generic sentence grammar network. The references are not applicable. Notice that nowhere in Naylor does it mention multiple sets and mapping of each of the sets independently to a generic grammar.

Claim 7 is dependent on Claim 6 and is therefore deemed allowable for at least the same reason as claim 6.

The examiner's rejection of Claim 8 states that:

"As per claim 8, Naylor et al. teach the method of claim 7, wherein said path propagation includes getting offset HMMs, offset symbols and the base symbol for a given expanded symbol and obtaining the HMM of the previous frame and expanding and storing a sequence set of HMM states both for within model path and cross model path and determining the path with the best transition probability (as using labels from stored data, variance, frequency of occurrence-(Col. 5, lines 55-65; wherein this data is merged with the original HMM data to formulate the new probabilities-Col.7, lines 40-55; Fig. 6)."

All HMM recognizers using a Viterbi search contain the processes of extending likelihood calculations for both within model paths and cross model paths and determining the paths with best transition probability. However, just as in claim 7, Naylor does not mention obtaining offsets that index each HMM set and retrieving the individual symbols for each HMM set that correspond to the base symbol within the generic sentence grammar, then extending the Viterbi search for each symbol for each HMM set individually and separately. This is, of course, because Naylor has nothing to do with a recognizer that handles a plurality of HMM sets and returns the result providing the best likelihood from a single HMM set.

The examiner in the rejection of Claim 9 does correctly state that an updated observation probability is included in the HMM processing, which is common to all Viterbi based HMM recognizers. However, the present application teaches that updating of observation probabilities occurs independently within each HMM set, so that we keep track of likelihoods individually with each HMM set. Naylor does not teach this, since his patent does not have anything to do with simultaneously recognizing separately multiple HMM sets.

Applicants newly added claim 10-19 are deemed allowable over the references for the reasons discussed above. Claim 12 calls for " means for constructing recognition paths defined on expanded-symbols wherein each expanded-symbol references a model contained in one of the model sets, and means for determining expanded-symbols by a conversion function that maps a base-symbol of the generic base grammar network to a plurality of expanded-symbols and an expanded-symbol to its corresponding base-symbol.." As discussed above this is neither taught not suggested in the references. Claims 13-19 dependent on claim 12 are deemed allowable for at least the same reasons as claim 12.

In view of the above applicants Claims 3, 5-8, as amended, and new claims 10-19 are deemed allowable and an early notice of allowance of these claims is deemed in order and is respectfully requested.

Respectfully requested;

Robert L. Troike

Robert L. Troike (Reg. 24183)

Telephone No.(301) 259-2089